



Co-funded by
the European Union



Deliverable D6.2: Data Management Plan

Deliverable D6.2

Contractual Date: 31-02-2025

Actual Date: 27-03-2025

Grant Agreement
No.: 101190015

Work Package: WP6

Task Item: T6.2

Lead Partner: EITD

Authors: Vilma Djala(EITD)

Abstract

This Data Management Plan (DMP) outlines how data within the project is managed in a FAIR (findable, accessible, interoperable, and reusable) manner. The DMP has been created in accordance with the Horizon Europe Data Management Plan Template. The DMP should include information on the handling of research data during and after the end of the project, what data will be collected, processed, and/or generated, how data will be curated and preserved, and other relevant issues. By following the guidelines outlined in the DMP, project partners can ensure that data is managed in a way that maximizes its value and impact.

© EIT Digital on behalf of the ACHIEVE project.

The activities leading to these results has received funding from the European Community's DIGITAL Programme under Grant Agreement No. 101190015 (ACHIEVE).

This document is licensed under a [Creative Commons Attribution 4.0 license](https://creativecommons.org/licenses/by/4.0/).





Co-funded by
the European Union



Versioning and contribution history

Version	Date	Authors	Notes
0.1	13/02/2025	Vilma Djala (EITD)	First version of D6.2
0.2	19/03/2025	Vilma Djala (EITD)	Second version of D6.2, with integrated comments from partners.
0.3	27/03/2025	Vilma Djala (EITD)	Final version of D6.2

© EIT Digital on behalf of the ACHIEVE project.

The activities leading to these results has received funding from the European Community's DIGITAL Programme under Grant Agreement No. 101190015 (ACHIEVE).

This document is licensed under a [Creative Commons Attribution 4.0 license](https://creativecommons.org/licenses/by/4.0/).



Table of Contents

1. Introduction.....	3
1.1 ACHIEVE.....	4
2. Work Package 6.....	5
3. Deliverable 6.2	5
3.1 Purpose.....	5
3.2 Objectives	5
4. Data summary.....	6
5. FAIR Data	11
5.1 Making data findable, including provisions for metadata	11
5.2 Making data accessible	11
5.3 Making data interoperable	13
5.4 Increase data re-use (through clarifying licences)	14
6. Allocation of resources.....	14
7. Data security	14
8. Ethical aspects.....	15
9. Other issues	16

Table of Tables

Table 1: Data sets overview	6
Table 2: Data sets description and utility	7
Table 3: Data sets accessibility	11

1. Introduction

This Data Management Plan (DMP) has been created following the Horizon 2020 FAIR DMP template, which has been designed to be applicable to any Horizon 2020 project that produces, collects or processes research data.

This document is intended to follow the best practices for a FAIR data management¹.

Definition: FAIR data management

In general terms, your research data should be 'FAIR', that is findable, accessible, interoperable and re-usable. These principles precede implementation choices and do not necessarily suggest any specific technology, standard, or implementation-solution.

This Data Management Plan is a set of questions, from the Horizon 2020 template, that were answered with a level of detail appropriate to the project. This DMP is intended to be a living document in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur.

This DMP, while following the Horizon 2020 FAIR template, reflects the specifics of the ACHIEVE project in several ways:

Data Types: The DMP explicitly categorizes the data collected into "non-sensitive" (publicly available) and "operational" (including sensitive data about students). This reflects the dual nature of data generated in educational projects, where both publicly releasable research data and private student information are

¹ FAIR Data Principles (FORCE11 discussion forum):

<https://force11.org/group/fairgroup/fairprinciples>

FAIR principles (article in Nature): <https://www.nature.com/articles/sdata201618>

D6.2 Data Management Plan

handled. A generic template wouldn't necessarily have this specific distinction built-in.

Data Accessibility: The DMP clearly outlines which data sets will be openly available and which will remain restricted due to privacy concerns (GDPR compliance). This detailed accessibility matrix is tailored to the specific data generated by ACHIEVE, and a generic template would offer a more general framework.

Project Specifics: The DMP directly addresses the needs and objectives of the ACHIEVE project, such as supporting the development of cloud and networking infrastructure, and its involvement in education and training programs. The data collected directly relates to these project goals, making the DMP more contextually relevant than a generic FAIR template.

Ethical Considerations: The DMP highlights ethical aspects specific to the handling of sensitive student data, explicitly mentioning informed consent and data privacy measures which are crucial to the ACHIEVE project but might not be the focus of a more generally applicable FAIR template.

As a minimum, the DMP should be updated in the context of the periodic evaluation/assessment of the project. If there are no other periodic reviews envisaged within the grant agreement, an update needs to be made in time for the final review at the latest.

The Data Manager responsible will be Carolina Fernandez Ramon, Legal & Compliance Manager at EITD. She will be reachable for any concern around data handling at carolina.fernandez@eitdigital.eu.

1.1 ACHIEVE

Advanced Cloud and High-performance computing Education for a Valiant Europe (ACHIEVE) aims at:

- Supporting the excellence of EU higher education strategic development of Cloud and Networking Infrastructure to obtain advanced, green, and efficient HPC systems, updating the academic curricula on this cutting-edge topic through the matching between the industry needs and the academic offer,
- Increasing the awareness of the impact of HPC systems in the society and citizens' daily life, and changing the perception of low exciting discipline,
- Providing startups and SMEs the opportunity of becoming hosts of students' internships.

D6.2 Data Management Plan

2. Work Package 6

The objectives of Work Package 6 are as follows:

- To oversee project management and provide effective monitoring across administrative, technical, and financial aspects.
- To ensure high-quality content and management practices that drive meaningful project progress.
- To coordinate the participant enrollment process for the ACHIEVE education programs.
- To foster the creation of effective and sustainable partnerships within the project consortium.

This work package focuses on leading the technical and scientific coordination of the ACHIEVE project, along with its administrative and financial management. It will also implement quality control and reporting mechanisms throughout the project.

3. Deliverable 6.2

3.1 Purpose

The ACHIEVE Data Management Plan has been prepared with two purposes:

1. to describe the data management life cycle for the data to be collected, processed and/or generated by the ACHIEVE project;
2. include information on the handling of research data during and after the end of the project, what data will be collected, processed and/or generated, how data will be curated and preserved, and resource and budgetary planning for data management.

3.2 Objectives

- Ensure effective management of research data throughout the project life cycle.
- Describe the data management life cycle for the data to be collected, processed and/or generated by the project.
- Ensure that research data is findable, accessible, interoperable and re-usable (FAIR).

D6.2 Data Management Plan

- Ensure that research data is managed in compliance with the General Data Protection Regulation (GDPR).
- Reflect the current state of consortium agreements on data management and be consistent with exploitation and Intellectual Property Rights (IPR) requirements.
- Provide an overview of all datasets collected and generated by the project and define the consortium's data management policy and approach.

4. Data summary

In order to provide an overview of the different data sets that are currently and will be produced in the ACHIEVE project, we need to distinguish two types of data:

1. Non-sensitive data produced by the project, publicly available, and released for potential reuse in other projects or research activities.
2. Operational data used to implement the activities described in the project. This data includes very frequently also sensitive data about students and participants to training activities.

In this plan, there is a consistent use of terms such as “learners”, “participants”, “students” and “candidates” almost interchangeably. Students and candidates refer to people who will be engaged in the application process and the learning within the Master Programme. Instead, terms like learners and participants refer to the self learning modules and/or to corollary activities that help build the learning programmes.

The following table shows the data type, the origin of the data, the related WP number and the format, in which the data will be presumably stored.

Table 1: Data sets overview

#	Data type	Type	Origin	WP#	Format
1	Market review of Cloud Computing and HPC	Non-sensitive	Derived data by other reports and market data.	WP1, WP2	.pdf
2	Literature review data on Cloud Computing and HPC	Non-sensitive	Derived data by publications or published reports.	WP1, WP2	.pdf

D6.2 Data Management Plan

#	Data type	Type	Origin	WP#	Format
3	Recruitment cycle data about participants.	Operational	Primary data	WP1, WP2	.xlsx and .docx and .pdf
4	Personal data of students participating to master programs.	Operational	Primary data	WP1, WP2	.xlsx and .pdf and university systems
5	Data on participants to self-standing modules.	Operational	Primary data	WP1, WP2	.xlsx and .pdf
6	Satisfaction survey from students at the end of a learning course or activity.	Operational	Primary data	WP1, WP2	.xlsx and .pdf
7	Marketing data related to communication and dissemination activities.	Operational	Primary data	WP5	.xlsx and .pdf

While the next table describes the data set and the purpose of the data collection of data generation in relation with the objectives of the project. Additionally, it shows the data utility for clarifying to whom the data might be useful.

Table 2: Data sets description and utility

1	Market review of Cloud Computing and HPC	<p>Description The data contains the result of a market review analysis done on the field of Cloud Computing and HPC. The analysis will be performed by analyzing publicly available market data and by interviewing economic actors in the sector.</p> <p>Purpose The collection of this data will serve as an</p>	<p>The data could be useful for research on the Cloud Computing and HPC sectors.</p> <p>It can also be useful for other educational institutions and to organizations and business to better understand the current state of the market,</p>

D6.2 Data Management Plan

		input to the process of review of the master program curriculum. This data will also serve to guide the definition of the content for the self-standing learning modules.	identify the latest trends and threats and make informed decisions about Cloud Computing and HPC products and services.
2	Literature review data on Cloud Computing and HPC.	<p>Description The data contains the result of a literature review done on the field of Cloud Computing and HPC. The analysis will be performed by analyzing publications, articles and course syllabi from other universities and higher education institutions.</p> <p>Purpose The collection of this data will serve as an input to the process of review of the master program curriculum on Cloud Computing and HPC. This data will also serve to guide the definition of the content for the self-standing learning modules.</p>	<p>The data could be helpful to researcher interested in understanding the current state of knowledge in the field of Cloud Computing and HPC, identify gaps in the literature and develop research question and hypotheses.</p> <p>The data can also be useful for policymakers by helping the development of policies and regulations that are evidence-based and effective.</p>
3	Recruitment cycle data about participants.	<p>Description This data includes all the personal information of candidates applying for the master programmes. The data will include contact information, CV history and study track records for all applicants.</p> <p>Purpose Data is gathered for administrative purposes and to enable the selection</p>	<p>Researchers can use this data, after anonymization, to study the qualifications and backgrounds of candidates applying for master's programs or jobs. The data can help researchers identify trends and patterns in the qualifications and backgrounds of successful candidates and develop</p>

D6.2 Data Management Plan

		of candidates, including the awarding of scholarships, based on their recent educational and professional history.	research questions and hypotheses. Personal data could be shared with relevant third parties (i.e. employers or recruitment agencies) upon collection of individual and informed consent of participants.
4	Personal data of students participating to master programs.	<p>Description This data includes all the personal information of students of the master programs. The data will include contact information, CV history and study track records for all students and will be managed by the guesting universities following the general rules for all students.</p> <p>Purpose Data is gathered for administrative purposes (i.e. admission process), to enable participation to the courses and track of the student's academic pathway .</p>	<p>Researchers can use this data, after anonymization, to study the qualifications and backgrounds of learners of the self-standing modules. The data can help researchers identify trends and patterns in the qualifications and backgrounds of online students and develop research questions and hypotheses.</p> <p>Personal data could be shared with relevant third parties (i.e. employers or recruitment agencies) upon collection of individual and informed consent of students.</p>
5	Data on participants to self-standing modules.	<p>Description Data related to the registration and participation to self-standing modules. This data includes contact information and digital addresses of all participants. The data also includes information regarding eventual certifications obtained by the participants.</p>	<p>Researchers can use this data, after anonymization, to study the qualifications and backgrounds of learners of the self-standing modules. The data can help researchers identify trends and patterns in the qualifications and backgrounds of online students and develop</p>

D6.2 Data Management Plan

		<p>Purpose Data is gathered for administrative purposes to enable the access to the online platform and the tracking of the study activities.</p>	<p>research questions and hypotheses. Personal data could be shared with relevant third parties (i.e. employers or recruitment agencies) upon collection of individual and informed consent of participants.</p>
6	Satisfaction survey from students at the end of a learning course or activity.	<p>Description Data related to the results of the satisfaction survey gathered from students of master and self-standing modules.</p> <p>Purpose Data is gathered to implement a quality improvement process and to improve courses and training material.</p>	<p>The data is of interest to the project participants to obtain workable feedback and encourage continuous improvement cycle of the courses materials and trainings paths. It will not be publicly available unless anonymized.</p>
7	Marketing data related to communication and dissemination activities.	<p>Description Data regarding the communication and dissemination campaign on social networks and digital channels.</p> <p>Purpose Digital marketing will be a central part of the strategy of attraction to candidate students to the project's program. Collecting operational data is fundamental for digital communication to work effectively.</p>	<p>The anonymized data could be helpful to digital marketing agencies or marketing professional interested in evaluating the effectiveness of their digital marketing campaigns and identify areas for improvement. Researchers can use this anonymized data to study the effectiveness of digital marketing campaigns and identify trends and patterns in the participation of potential students or customers.</p>

5. FAIR Data

5.1 Making data findable, including provisions for metadata

To ensure that the data generated during the project is findable, we will implement the following provisions:

- All data will be recorded in a predeterminate structure and with agreed data formats.
- Data structure and format will ensure interoperability and ease of use.
- To ensure that the data is discoverable, we will implement the following mechanisms:
 - Data will be assigned unique identifiers to enable easy identification and tracking.
 - Data will be stored in a structured and organized manner to enable efficient searching and browsing.
 - Data will be made available through appropriate repositories and archives to enable discovery and reuse.

5.2 Making data accessible

The following table is highlighting which data is intended for internal use (within ACHIEVE project) and which will be made openly available. It also explains why several datasets cannot be shared because of particular reasons and, in this case, an alternate solution will be presented.

Table 3: Data sets accessibility

#	Data type	Openly available	Justification	Alternate solution
1	Market review of Cloud Computing and HPC sectors.	Yes	Results of this analysis will be described in the project deliverable D1.1.	(not relevant)

D6.2 Data Management Plan

#	Data type	Openly available	Justification	Alternate solution
2	Literature review data on Cloud Computing and HPC	Yes	Results of this review will be described in the project deliverable D1.1.	(not relevant)
3	Recruitment cycle data about participants.	No	The sensitive data about students involved in the recruitment process of the project will not be released, in respect to GDPR and any other regulation that may apply.	Statistical data about the recruitment cycle and student admissions will be described in the project deliverables D4.4, D4.5, D4.6, D4.7.
4	Data of students participating to master programs.	No	The sensitive data about students participating to courses and learning activities of the project will not be released, in respect to GDPR and any other regulation that may apply.	Statistical data about the student participation to the master will be described in the project deliverables D1.2, D1.3, D1.4.
5	Data on participants to self-standing modules.	No	All personal data, including contact information, regarding students, will not be made openly available, in respect to GDPR and any other regulation that may apply.	Statistical data about the student participation to the master will be described in the project deliverables D3.2, D3.3.
6	Satisfaction survey from students at the end of a learning	No	All personal data of students, including contact information and opinion on the course quality, will not be made	Statistical data about the student satisfaction expressed for the courses attended will be

D6.2 Data Management Plan

#	Data type	Openly available	Justification	Alternate solution
	course or activity.		openly available, in respect to GDPR and any other regulation that may apply.	described in the project deliverables D1.2, D1.3, D1.4, D2.2, D2.3.
7	Marketing data related to communication and dissemination activities.	No	The granular and analytical marketing data used to guide communication and dissemination activities will not be released.	Statistical aggregated data about marketing and dissemination activities will be released in deliverables D5.2, D5.3, D5.4, D5.5.

The data intended for open availability, will be registered in official project deliverables and, as such, will be published on the project website and on the EC portal for public access.

5.3 Making data interoperable

All the data shared by the project will use document standards that will make it interoperable. The aim is to achieve a standardized, open, and flexible way to exchange and reuse data across different systems and applications.

5.4 Increase data re-use (through clarifying licences)

To permit the widest re-use of data, all openly available project deliverables and main results will be released with a Creative Commons Attribution (CC-BY) license. This license allows others to distribute, remix, and build upon the data, even commercially, as long as they credit the original source.

The data released under this license does not include:

- any sensitive information regarding students, that will be protected by GDPR and other relevant regulation,
- the master's course materials, that will remain property of the producing entity, and
- the online training modules, that will maintain a shared ownership between the beneficiaries that have generated them.

6. Allocation of resources

The following resources will be allocated to ensure effective data management throughout the project:

- **Personnel:** Data management will be overseen in Task 6.5 of the WP6. The task will permit all partners involved in these activities to dedicate resources, including personnel, to the tasks and activities related to data management. A data manager will be appointed to oversee the implementation of the data management plan and ensure compliance with relevant regulations and guidelines. The data manager will be responsible for creating and maintaining the metadata, ensuring data quality, and managing the storage and security of the data. The data manager will also be responsible for training project personnel in data management best practices.
- **Infrastructure:** The project will allocate resources for the storage and backup of data in secure locations. For this purpose, the project will use the website and the Teams instance of the coordinator partner: EIT Digital.
- **Budget:** The project will allocate a budget for data management activities. The budget will also include provisions for the dissemination and sharing of data, including the use of appropriate repositories and archives.
- **Deliverables:** This deliverable D6.2 of the project describes an initial data management plan. The project will also include deliverables for the dissemination and sharing of data.

The allocation of resources will be reviewed and updated throughout the project as necessary to ensure that the data management plan remains effective and compliant with relevant regulations and guidelines. The project will also ensure that the allocation of resources is consistent with exploitation and Intellectual Property Rights (IPR) requirements.

7. Data security

The project will implement the following measures to ensure the security of the data recorded on the Microsoft Teams platform used to store all relevant project data:

8. **Access controls:** Access to the data will be restricted to authorized personnel only. The project will use Microsoft Teams to manage access

D6.2 Data Management Plan

controls, including role-based access controls and multi-factor authentication.

9. **Backup and recovery:** The project will implement a backup and recovery plan to ensure that the data is recoverable in the event of a disaster or system failure. The project will use the Microsoft's platform backup and recovery capabilities to ensure that the data is protected.
10. **Data retention and disposal:** The project will implement a data retention and disposal policy to ensure that data is retained only for as long as necessary and disposed of securely when it is no longer needed.
11. **Monitoring and auditing:** The project will implement monitoring and auditing procedures to ensure that the data is being used appropriately and that any unauthorized access or use is detected and addressed.

The project will also ensure that all personnel involved in the project are trained in data security best practices and that they understand their roles and responsibilities in protecting the data. The project will also ensure that all data management activities are compliant with relevant regulations and guidelines.

8. Ethical aspects

The project will ensure that all data management activities are conducted in compliance with relevant ethical guidelines and regulations. The following ethical aspects will be considered:

1. **Informed consent:** The project will obtain informed consent from all participants before collecting any data. Participants will be informed about the purpose of the data collection, how the data will be used, and any potential risks or benefits associated with the data collection.
2. **Data privacy:** The project will ensure that all data is collected, stored, and shared in compliance with relevant data privacy regulations. The project will implement appropriate measures to protect the privacy and confidentiality of the data, including encryption, access controls, and data anonymization where necessary.
3. **Data ownership:** The project will ensure that all data is owned by the appropriate parties and that any intellectual property rights are respected. The project will also ensure that any data sharing or dissemination is conducted in compliance with relevant regulations and guidelines.
4. **Data sharing:** The project will ensure that any data sharing or dissemination is conducted in compliance with relevant regulations and

D6.2 Data Management Plan

guidelines. The project will also ensure that any data sharing or dissemination is conducted in a manner that respects the privacy and confidentiality of the data.

5. **Data retention and disposal:** The project will implement a data retention and disposal policy to ensure that data is retained only for as long as necessary and disposed of securely when it is no longer needed.

9. Other issues

Data management in the project will be performed following the European Commission's Horizon 2020 procedures. In particular this document represents the Data Management Plan (DMP) as requested in the program and describes the data management life cycle for the data to be collected, processed, and/or generated by the project.

D6.2 Data Management Plan

References

[DIGITAL] <https://digital-strategy.ec.europa.eu/en/activities/digital-programme>

[ACHIEVE] <https://eitictlabs.sharepoint.com/sites/ACHIEVEProjectManagement>

Glossary

Community

A group of users, organised with a common purpose, and jointly granted access to resources. It may act as the interface between individual users and the resources. (see also [[WISE-SCI](#)])

DMP	Data Management Plan
EC	European Commission
EIT	European Institute of Innovation and Technology
KIC	Knowledge and Innovation Community
GA	Grant Agreement
GDPR	General Data Protection Regulation
HPC	High Performance Computing
R&S	Research and scholarship
SME	Small and Medium Enterprise